

<https://helda.helsinki.fi>

Digitaalsete keeleressursside loomist ja kasutamist määrav õiguslik raamistik Eestis ja selle ühildumine CLARIN-i infrastruktuuriga

Kelli, Aleksei

2016

Kelli , A , Vider , K , Pisuke , H & Linden , B K J 2016 , ' Digitaalsete keeleressursside loomist ja kasutamist määrav õiguslik raamistik Eestis ja selle ühildumine CLARIN-i infrastruktuuriga ' , Eesti rakenduslingvistika ühingu aastaraamat , vol. 12 , pp. 81-98 . <https://doi.org/10.5128/ERYa12.05>

<http://hdl.handle.net/10138/162994>

<https://doi.org/10.5128/ERYa12.05>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

DIGITAALSETE KEELERESSURSSIDE LOOMIST JA KASUTAMIST MÄÄRAV ÕIGUSLIK RAAMISTIK EESTIS JA SELLE ÜHILDUMINE CLARIN-i INFRASTRUKTUURIGA

Aleksei Kelli, Kadri Vider,
Heiki Pisuke, Krister Lindén

Ülevaade. Artikkel keskendub keeleressurssidega seonduvatele õiguslikele küsimustele. Aluseks võetakse protsess, mis algab keeleressursside loomisega ning lõpeb nende levitamisega. Keeleressurss on oma olemuselt autoriõiguslikult kaitstav andmebaas, mida katavad kahetasemelised õigused (õigused autoriõiguslikult kaitstavale materjalile, mida ressursside loomisel kasutati, ja õigused ressursile endale), millega tuleb levitamise protsessis arvestada.

Keeleressursside loomine saab tugineda kahele võimalikule õiguslikule mudelile: 1) litsentsipõhine mudel ja 2) erandipõhine mudel. Mõlemaga kaasnevad positiivsed ja negatiivsed aspektid, mida käesolevas artiklis ka analüüsitakse, võttes arvesse arenguid Euroopa Liidus.

Enne kui keeleressurssi saab levitada CLARIN-i võrgustikus, on oluline, et õigused levitada oleks üle antud mõnele Eesti Keeleressursside Keskuse konsortsiumipartnerile. Õigused ressursile lähevad üle kas töösuhtes või siis lepinguga. Sel eesmärgil on töötatud välja CLARIN-i deponeerimislepingud. Enne kui kasutaja saab ressursile ligipääsu, peab ta nõustuma CLARIN-i teenusetingimustega.

Artikkel selgitab CLARIN-i tüüplepingutega seonduvaid küsimusi ning tutvustab CLARIN-i lepingute viimaseid parandusi.

Võtmesõnad: keeleressursside õiguslik ja lepinguline raamistik, autoriõigus, keeleressursside kategooriad, CLARIN-i tüüplepingud, vastutus

1. Sissejuhatus

Keel ei ole ammu enam üksnes suhtlemisvahend, teaduslik uurimisobjekt ja inimeste identiteedi oluline osa. 21. sajandil omab üha suuremat tähtsust keel kui informatsioon. Selle informatsiooni kogumisel on mitmed erinevad eesmärgid ja meetodid. Samuti on sellisel informatsiooni iseseisev turuväärtus. Teatud viisil kogutud ja korrastatud digitaalset keelelist informatsiooni nimetatakse keeleressurssideks (ingl *language resources*).

Keeleressursside loomine toimub riigisiselt, teatud riikide gruppide poolt või Euroopa Liidu (EL) tasandil.¹ Kõikidel nendel tasemetel on oluline roll õiguslikel küsimustel, sealhulgas autoriõigusel.

Käesolevas artiklis on põhitähelepanu CLARIN-i võrgustiku² (CLARIN) ja Eesti koostööl seoses keeleressursside puudutavate õiguslike küsimustega.

CLARIN-i üks eesmärk on koordineerida oma liikmete tegevust ja ühtlustada (standardiseerida) teatud valdkondi. See hõlmab ka õiguslikku reguleerimist. Ideaalis peaks keeleressursside ja -tehnoloogiate (edaspidi koos ka keeleressurssid) loomise ja levitamise õiguslik keskkond olema sarnane kõigis CLARIN-i liikmesriikides. Sisuliselt tähendab see õiguslike reforme EL-i tasandil, kuna paljud olulised autoriõiguse küsimused on EL-i pädevuses. Kuigi CLARIN on astunud samme selle eesmärgi suunas (osalenud erinevates sidusgruppide konsultatsioonides EL-i ja siseriiklikul tasandil ning andnud sisendit EL-i poliitikatesse), jääb keeleressursside ja -tehnoloogiaid ümbritseva õigusliku raamistiku harmoniseerimine väljakutseks, millega tuleb võrgustiku liikmetel jätkuvalt tegeleda.

Käesoleva artikli eesmärk on uurida keeleressursside loomise ja levitamise seonduvaid õiguslikke küsimusi ja riske. Taoline analüüs aitab keeleressursside loojatel, kasutajatel ja ka teistel sidusgruppidel (näiteks keeleressursside loomise finantseerijatel ja äriühingutel, kes on huvitatud keeleressursside ja -tehnoloogiate ärilisest kasutamisest) oma tegevust paremini mõtestada. Artikli rõhuasetus on keeleressurssidega seotud lepingulisel raamistikul (CLARIN-i tüüplepingutel). Kuna lepingulise raamistiku küsimused on lahutamatu seotud keeleressursside loomisega, siis käsitlevad autorid ka keeleressursside loomisega seonduvat õiguslikku problemaatikat (keeleressursside loomise autoriõiguslike aspektide kohta vt Kelli jt 2012, Tavast jt 2013).

Lisaks autoriõigusele ja kaasnevatele õigustele on keeleressursside kontekstis aktuaalne ka isikuandmete kaitse, millel artiklis pikalt ei peatuta. Isikuandmete kaitse teemal keeleressursside kontekstis on avaldatud uurimus, mille seisukohtadest käesolevas artiklis lähtutakse (vt Jents, Kelli 2014).

Artikkel tugineb autorite praktilisele kogemusele, milleks on aastatepikkune tegelemine autoriõiguse, keeleressursside autoriõiguslike probleemidega, osalemine CLARIN-i õigusasjade komitees ja CLARIN ERIC-u riiklik koordineerimine Eestis ning Eesti Keeleressursside Keskuse juhtimine. Enne kui võtta konkreetsemalt vaatluse alla õigusliku raamistikuga seonduvad probleemid, peatuvad autorid keeleressursside ja -tehnoloogia õiguslikel määratlustel.

Keeletehnoloogia vajab arendamiseks ja toimimiseks üsna mitmekesiseid keeleressursside: arvutiprogramme, muid teoseid (leksikonid, grammatikad, korpused),

¹ Käesolevas artiklis ei käsitleta uusimat Euroopa Komisjoni initsiatiivi EL-i mitmekeelsuse poliitika raames: 2015. a käivitunud European Language Resource Coordination projekti ja European Language Resource Coordination Consortiumi (ELRC) tegevust, milles osaleb 30 Euroopa riiki, sh kõik EL-i liikmesriigid, Island ja Norra.

² CLARIN-i liikmeteks on Austria, Bulgaaria, Tšehhi Vabariik, Taani, Hollandi Keele Ühendus (*Dutch Language Union*), Eesti, Saksamaa, Kreeka, Leedu, Holland, Norra, Poola, Portugal, Sloveenia, Rootsi, Soome, Ühendkuningriik.

andmebaase jms. Autoriõiguse seadus (AutÕS) kaitseb ka arvutiprogramme; arvutiprogramme kaitstakse sama õigusliku režiimi alusel nagu kirjandusteoseid (§ 4).

Autoriõiguse seadus reguleerib kahte liiki iseseisva tähendusega õigusi, mille maht ja kaitse tähtajad on erinevad. Need on autoriõigused ja autoriõigusega kaasnevad õigused. Loodud keeleressursid on meie arvates enamuses autoriõiguslikult kaitstavad teosed. Näiteks keelekorpust võib käsitleda andmebaasina, mis sisaldab kirjalikke ja suulisi tekste. Autoriõiguse seaduse kohaselt peab teos olema originaalne (AutÕS § 4). Originaalsus väljendub autori enda loomingus, mis eeldab valikuvabadust. Keeleressursside kontekstis väljendub originaalsus tekstide valimises, mis on loominguline tegevus. Kõrvuti autoriõigusega on keeleressursid kaitstavad ka autoriõigusega kaasnevate õigustega; seaduses on eraldi peatükk nn mitteloominguliste andmebaaside kohta, kus kaitse objektiks on peamiselt sellesse andmebaasi tehtud investeering. Erialakirjanduses nimetatakse selliseid autoriõigusest eraldi ja nende kõrval eksisteerivaid õigusi *sui generis* õigusteks, et eristada neid autoriõigusest.³ Käesolevas artiklis nimetatakse selliseid andmebaase *sui generis* andmebaasideks. Autoriõiguse seaduse kohaselt on *sui generis* andmebaas “teoste, andmete või muu materjali süstemaatiliselt või metoodiliselt korraldatud kogu, mis on individuaalselt kasutatav elektrooniliste või muude vahendite abil. Andmebaasi mõiste ei hõlma selle tegemiseks ega käivitamiseks vajaminevat arvutiprogrammi” (AutÕS § 75²). Õigused *sui generis* andmebaasile kehtivad 15 aastat, kuid andmebaasi muutmisel pikeneb kaitse 15 aastale (AutÕS § 75⁷). Oluline nüanss siinjuures on see, et kaitse saadakse just muudatustele, mitte kogu andmebaasile. Kuna autoriõigusliku andmebaasi kaitse on tähtajalt pikem ja ulatuslikum, siis *sui generis* kaitse antud kontekstis olulist tähendust ei oma.

Artikli esimeses osas peatutakse keeleressursside loomisega seonduval õiguslikul teematikal. See on lahutamatu seotud keeleressursside lepingulise raamistikuga. Lisaks CLARIN-i tüüplepingute analüüsile selgitatakse ka nende täiendamise vajadust. Artikkel lõpeb soovitustega poliitikakujundajatele, keeleressursside loojatele ja kasutajatele.

2. Keeleressursside loomise võimalused autoriõiguse perspektiivist

Keeleressursside õiguslike küsimustega tegelemiseks tuleb analüüsida nende loomisprotsessi. Keeleressursid koosnevad elava keele materjalist, mida loovad kõik keele kasutajad ehk kõik inimesed. Põhimõtteliselt võib inimese iga keeleline eneseväljendus olla looming ja inimene selle loomingu autor. Keeleressursside loomiseks kasutatav keelematerjal võib pärineda raamatust, ajakirjast, raadiosaatest, blogist, netifoorumist, kommentaariumist või muust allikast. Kuna looming on kaitstav teosena, siis õiguslikud küsimused seonduvad autoriõigusega. Kui kasutatud teksti autori surmast on möödunud 70 aastat (teos on kaitstav 70 aastat pärast autori surma) või on autoriõigusliku kaitse alt välistatud (rahvalooming, õigusaktid, kohtulahendid), siis on see vabalt kasutatav. Oluline on arvestada autoriõigusega kaasnevate õiguste regulatsiooniga. Autoriõiguse seadus annab varem avaldamata teose avaldajale autori varaliste õigustega sarnased õigused 25 aastaks. Autoriõiguse

³ Autoriõiguse seaduses on *sui generis* andmebaasi regulatsioon VIII¹ peatükis (andmebaasi tegija õigused).

seaduse kohaselt (§ 74¹) on isikul, kes pärast autoriõiguse kehtivuse tähtaja lõppemist esimesena õiguspäraselt avaldab või suunab üldsusele varem avaldamata teose, 25 aasta jooksul sellise teose esmakordsest avaldamisest või üldsusele suunamisest autori varaliste õigustega võrdsed õigused sellele teosele. See tähendab, et ka ajaloolise teksti kasutamisega võib rikkuda kellegi kaasnevaid õigusi.

Suulise (mõnikord ka kirjaliku) teksti kasutamisele lisanduvad isikuandmete kaitse ja isikuõiguse küsimused. Suulise ja audiovisuaalse teose puhul tuleb arvestada ka autoriõigusega kaasnevate õiguste regulatsiooniga. Kaasnevad õigused kaitsevad esitaja⁴ õigusi, aga samuti fonogrammitootja ja ringhäälinguorganisatsiooni õigusi. Ei saa välistada, et tekst lisatakse andmebaasi fonogrammilt, raadio- või telesaatest. Isikuandmeid sisaldav tekst allub ka isikuandmete kaitse regulatsioonile. Isikuandmete kaitse seadus (IKS) määratleb isikuandmeid järgmiselt: "Isikuandmed on mis tahes andmed tuvastatud või tuvastatava füüsilise isiku kohta, sõltumata sellest, millisel kujul või millises vormis need andmed on" (§ 4). Lisaks näeb võlaõigusseadus (VÕS) ette kaitse isikuõigustele ja isiklikele õigustele. Võlaõigusseadus käsitleb isikuõiguse (ka isiklike õiguste) riivena isiku nime või kujutise õigustamatut kasutamist ja eraelu puutumatuse rikkumist (§ 134, 1046). Selle teemaga seonduvad ka autori ja esitaja isiklike õiguste kaitse küsimused. Kuna isiklike õiguste teostamise teemat on autorid eelnevalt komplekselt analüüsinud (Kelli jt 2014), siis siinkohal piirduetakse järeldusega, et nõusolek isikuandmete kasutamiseks ja isikuõiguste teostamiseks ei tohiks oma olemuselt erineda. Samas tuleb rõhutada, et nii isikuandmete kaitse kui ka isikuõiguste õiguslik režiim on mõlemad iseseisvad.

Niisiis võib ütelda, et keeleressursside loomine on seotud autoriõiguse, autoriõigusega kaasnevate õiguste ning isikuandmete kaitse ja isikuõiguste õiguslike režiimidega.

2.1. Keeleressursside loomise mudelid

Autoriõiguslikust perspektiivist saab keeleressursside loomine tugineda kahele erinevale mudelile: nõusolekumudelile ehk litsentsipõhisele mudelile ja erandipõhisele mudelile. Mõlemal mudelil on nii poolt- kui ka vastuargumente.

1) Nõusolekumudel ehk litsentsipõhine mudel. Nõusolekul põhinev mudel tähendab seda, et küsitakse nõusolek autoriõiguse, autoriõigusega kaasnevate õiguste omajalt (nt esitajalt) ja andmesubjektilt.

Selle mudeli positiivne külg on, et see arvestab kõige enam autoriõiguste ja kaasnevate õiguste omaja õigusi. Puuduseks on keeleressursside loomisega kaasnev suur administratiivne koormus (tehingukulu), sest eksisteerib vajadus 1) tuvastada isikud (autorid või autoriõiguste omajad), kellelt on vaja nõusolekut küsida; 2) selgitada neile, milleks nõusolekut vajatakse; 3) investeerida palju ressursi kogu süsteemi loomiseks ja haldamiseks. Mõne keeleressursi loomine sellistes tingimustes ei pruugi üldse võimalikuks osutuda, sest alati ei õnnestu autorit tuvastada (näiteks anonüümsed postitused, blogid), autorid ei anna nõusolekut (näiteks põhjusel, et nad ei saa keeleressursside loomise mõttest ja vajadusest aru) või on õigustatud isikute arv nii suur, et puudub administratiivne suutlikkus neilt kõigilt nõusolekuid saada (keeleressurss kui keeletehnoloogiale vajalik andmestik eeldab väga suurt sisendit ja on võrreldav massdigiteerimise projektidega).

2) Erandipõhine mudel. Erandil põhinev mudel tähendab seda, et seadus piirab autoriõiguste omaja või andmesubjekti (isiku, kelle kohta kasutatavad andmed käivad) õigusi. See on õiguste piirangul põhinev mudel.

Selle mudeli positiivne külg on võimalus vältida suurt tehingukulu, mis kaasneb nõusolekute küsimisega. Säästetud aja ja energia saab paigutada keeleressursside loomisse.⁵ Oluline on siinkohal ka teadvustada, et eesti keele digiteerimine ja digitaalsete keeleressursside loomine toimub avalikes huvides. Ilma keeleressurssideta jääks arendamata keeletehnoloogia, mida vajab digimaailmas ellujäämiseks ka põhiseadusega kaitstav eesti keel. Puudusena võib välja tuua Eesti seadusandja vabaduse piiratust seoses autoriõiguslike erandite ja piirangute kehtestamisega. Enamik piiranguid tuleneb mitmetest Euroopa Liidu direktiividest, eeskätt direktiivist autoriõiguse ja sellega kaasnevate õiguste teatavate aspektide ühtlustamise kohta infoühiskonnas (Infoühiskonna direktiiv).

Teatud tingimuste järgimisel on teose vaba kasutamise ehk autoriõiguslike erandite küsimus üks kõige päevakajalisemaid ja samas raskemaid teemasid Euroopa Komisjoni ja Euroopa Parlamendi päevakorras. EL-i tasandil toimub praegu intensiivne poliitiline ja teaduslik diskussioon erandite tuleviku üle, mis kahtlemata mõjutab ka Eestit.

2.2. Teksti- ja andmekaeve erandiga seotud arengud EL-is ja Eestis

Euroopa Komisjon viis 2013. a detsembri ja 2014. a märtsi vahel läbi ulatusliku konsultatsiooni kõigi EL-i huvigruppide hulgas, milles eraldi grupp küsimusi puudutasid erandite temaatikat (sh teksti- ja andmekaevet). Põhiliste probleemidena tõid huvigrupid teksti- ja andmekaeve juures välja õigusselguse puudumise ja probleemid litsentseerimise mehhanismiga. Keeleressursside loojad pooldasid autoriõigusliku erandi kehtestamist, mis võimaldaks keeleressursside loomist ilma nõusoleku küsimiseta, sest nõusoleku küsimisega kaasnevad barjäärina toimivad tehingukulud. Õiguste omajad (kirjastajad jt) on pigem autoriõigusliku erandi vastu ning toetavad litsentsipõhist lähenemist. Muuhulgas juhivad nad tähelepanu asjaolule, et autoriõiguslik erand ei lahendaks isikuandmete kaitsega seonduvaid probleeme. Õiguste omajaid esindavad organisatsioonid on seejuures pakkunud, et kui autoriõiguslik erand viiakse ikkagi sisse, siis sellega peaks kaasnema hüvitise maksmise kohustus (Avalik konsultatsioon 2014: 63–67).

Esimene suur teelahe tuleviku määramisel on vastus küsimusele, kas suudetakse kokku leppida autoriõiguse ja kaasnevate õiguste täielikus ühtlustamises EL-i ulatuses (seda nimetatakse EL-i dokumentides nn *single or unitary EU copyright title*). Selles suunas tegutseb nii Euroopa Komisjon kui ka rühm tuntud Euroopa intellektuaalse omandi professoreid, kes valmistasid ette Euroopa autoriõiguse mudelseaduse (Euroopa autoriõiguse koodeks 2010). Üleüldine autoriõiguse ühtlustamine vormis, mis asendaks rahvuslikud seadused, on intellektuaalselt huvitav ja kahtlemata edasise Euroopa integratsiooni huvides toimuv tegevus. Visioonina väärrib see toetust. Tuleb aga nõustuda nii poliitikute, mõjukamate huvigruppide kui ka paljude erialaspetsialistide hulgas enamlevinud seisukohaga, et praeguses EL-i arhitektuuris ja poliitilises olustikus ei ole selle idee elluviimine reaalne

⁵ Sidusgrupid on juhtinud siiski tähelepanu asjaolule, et kokkuleppe puudumisel õiguste omajaga, kellele kuulub autoriõiguslikult kaitstud sisu, võib selle kättesaadavus olla raskendatud tulenevalt tehnoloogilistest kaitsemeetmetest (Avalik konsultatsioon 2014: 66).

(Avalik konsultatsioon 2014: 89–91). Samas on teatud ühtlustamine hädavajalik, et saavutada EL-i ees seisvaid poliitilisi ja majanduslikke eesmärke. See on selgelt väljendatud Euroopa Komisjoni poliitikas, mida on asutud 2015. aastal ellu viima digitaalse ühtse turu (*Digital Single Market* ehk DSM) strateegia raames. DSM-i ühe eesmärgina on fikseeritud ühtlustatud erandite abil tagatav suurem õiguskindlus sisu piiriüleseks kasutamiseks konkreetsel otstarbel nagu teadustöö, haridus, teksti- ja andmekaeve (Euroopa digitaalse ühtse turu strateegia 2015, punkt 2.4). Euroopa Parlament on autoriõiguse tuleviku küsimuses samuti võtnud aktiivse positsiooni ja andnud komisjonile selged suunised oma eelistuste kohta erandite küsimuses (Ettepanek infoühiskonna direktiivi rakendamiseks).

Erandite küsimus on iseenesest poliitiline ja pakub mitmeid lahendusvõimalusi. EL-i intellektuaalse omandi poliitika majandusliku aluse analüüsimisel pööratakse üha rohkem tähelepanu otsese seose rõhutamisele intellektuaalse omandi ja majanduskasvu (sh tööhõivekasvu) vahel. Viimased uurimused näitavad, et autoriõiguse erandite paindlik režiim toetab majanduskasvu rohkem kui jäigalt fikseeritud erandite süsteem (Gibert 2015: 3–4). Õiguslikust seisukohast soosib selline lähenemine USA *fair use* doktriini (doktriini kohta vt Gorman, Ginsburg 2006: 715) ja EL-i maadest näib sellele kõige lähemal olevat Ühendkuningriigi *fair dealing* doktriin (doktriini kohta vt Cornish, Llewelyn 2007: 473–477). Samas on mandri-Euroopa traditsiooniline lähenemine nn otseselt kirjutatud erandite süsteem, mis on aluseks ka 2001. a infoühiskonna direktiivile. Ka Eesti autoriõigus on arenenud seaduses antud ammeldava erandite loetelu pinnal. Viimastel aastatel on aga ka Eesti õigusteaduses kerkinud USA *fair use*-iga analoogilise lähenemise otstarbekuse küsimus EL-is ja Eestis (vt Nemvalts jt 2012: 17).

Infoühiskonna direktiivis fikseeritud erandite endi tuleviku osas on üheks vaidlusküsimuseks see, kas erandid peavad olema kõigis liikmesriikides kohustuslikud, ja kui kõik erandid ei ole kohustuslikud, siis vähemalt need, mis on otseselt seotud siseturu tagamisega (Hohmann, Kullas 2015: 7–9). Intellektuaalomandi õiguste kaubandusaspektide lepingusse (TRIPS leping, art 13), Berni kirjandus- ja kunstiteoste kaitse konventsiooni (Berni konventsioon, art 9(2)) ja infoühiskonna direktiivi (art 5.5)⁶ sisse kirjutatud kolmeastmelisest testist (ingl *three-step test*) ei ole autorite ja loometööstuse tagapõhjaga huvigrupid nõus loobuma ka praeguse autoriõiguse reformi käigus. TRIPS lepingus, Berni konventsioonis ja infoühiskonna direktiivis sisalduv kolmeastmeline test on suunatud liikmesriikidele (Correa 2007: 134, Ricketson, Ginsburg 2006: 763, Walter, von Lewinski 2010: 1060). Kolmeastmeline test siseseviiduna siseriiklikku õigusesse (vt AutÕS § 17) on autori huvide kaitse garantii (Pisuke 2006: 79).

Viimase küsimusena tuleb mainida ka nn teadustöö erandi täiesti erinevat regulatsiooni eri liikmesriikides ja küsimust, kas see erand kvalifitseerub niinimetatud üle-Euroopalisele kohustuslikule, siseturu loomisele suunatud erandi staatusele. Liikmesriikide õigused uute erandite kehtestamiseks on infoühiskonna direktiiviga piiratud. Seega on Eesti jaoks oluline küsimus, kui kaugelt me saame Eestis minna nn teadustöö erandi tõlgendamisega, püüdes seda kohaldada ka keeleressurssidele.

Autoriõiguse seaduse kehtiv versioon (§ 18 ja 19) näeb ette mitmeid teose vaba kasutamise juhte, mis on asjakohased keeleressursside loomisel (teose isiklik kasutamine, teose tsiteerimine, teose kasutamine teaduslikel eesmärkidel, teose

⁶ 5.5. "...erandeid ja piiranguid kohaldatakse üksnes teatavatel erijuhtudel, mis ei ole vastuolus teose või muu objekti tavapärase kasutusega ning ei mõjuta põhjendamatult õiguste valdaja õiguspäraseid huve."

kopeerimine teaduslikel eesmärkidel). Autoriõiguse seadus (§ 75) näeb analoogsed piirangud ette ka kaasnevate õiguste suhtes. Samuti luuakse teatud osa eesti keeleressursse välismaal ja sellega seoses tuleb arvestada autoriõiguse seaduse ruumilist kehtivust.

Justiitsministeeriumi egiidi all läbi viidud intellektuaalse omandi kodifitseerimise ja reformi (2012–2014) eesmärgiks (intellektuaalse omandi reformi kohta vt Kelli 2015) oli muuhulgas muuta ka keeleressursside loomise regulatsioon selgemaks. Autoriõiguse ja autoriõigusega kaasnevate õiguste seaduse eelnõu (Autoriõiguse eelnõu) lubab õiguste eset⁷ kopeerida ja töödelda teksti- ja andme-kaeve eesmärkidel ning tingimusel, et selline kasutamine ei taotle ärilisi eesmärke. Oluline on lisada võimalusel ka viide algteosele või esitusele (§ 43 lg 1 punkt 3). Seletuskirjas autoriõiguse ja autoriõigusega kaasnevate õiguste seaduse eelnõu juurde (Autoriõiguse eelnõu seletuskiri) põhjendatakse seda järgmiselt:

“Keeleressursside, kultuuripärandi jm uurimisel, säilitamisel ning arendamisel on määrav tähtsus teoste ja esituste kasutamisel. Kehtiv autoriõiguse seadus sätestab teose kasutamisel teadustöö erandi ning selles raamistikus on keeleressursse, kultuuripärandit jm võimalik uurida ja arendada vaid teatud piirangutega. Muutmise ettepanekuga soovitakse teadustöö erandit täpsustada ja tagada õigusselgus. Selle eesmärgiks on tagada paremad võimalused vastavaks uurimis- ja arendustööks. Eesti keele digitaliseerimine on olulise tähtsusega eesti keele säilimise tagamisel”.

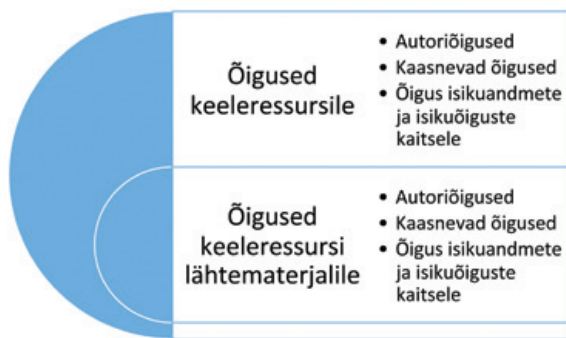
Keeleressursside arendajate ja ka riigi seisukohalt on oluline, et viidatud muudatused jõustuksid.

2.3. Keeleressursside kahetasemelised õigused

Keeleressursside loomise õiguslikult keerulist konstruktsiooni kirjeldab joonis 1. Esiteks tuginevad keeleressursid tihti teiste isikute varasemale loomingule. See tähendab, et keeleressursside loomiseks on vaja sisendmaterjali, mis on tihti kaetud autoriõiguste, kaasnevate õiguste ja ka isikuandmete ja isikuõiguste kaitsega. See moodustab n-ö keeleressursse katva õiguse esimese taseme. Teiseks tekivad õigused ka keeleressursside loojatel. Kuna keeleressurss on autoriõiguslikult kaitstav teos (näiteks andmebaas), siis selle loojal tekivad samuti õigused. See omakorda moodustab keeleressursse katva õiguse teise taseme.

Keeleressursse katvate õiguste kahetasemelisus (vt joonis 1) muutub aktuaalseks keeleressurssidega seonduvate tehingute (sh litsentsilepingute) kontekstis. Nimelt juhul, kui kolmandale isikule (näiteks Eesti Keeleressursside Keskusele) antakse keeleressursid üle või lubatakse neid kasutada. Ressursi looja võib oma õigused enda soovi kohaselt üle anda. Samas ei kuulu ressursi loojale keeleressursi loomiseks kasutatav keeleline lähtematerjal kui teos või teose osa. Kui keeleressursid on loodud tuginedes autoriõiguse seadusest tulenevale teose vaba kasutamise erandile teadustöö eesmärkidel (§ 19), siis saab ressursse kasutada samuti üksnes teadustööks. Järgnevas keskendumegi keeleressursside lepingulistele küsimustele.

⁷ Autoriõiguse eelnõu kohaselt (§ 2 (1)) on õiguste ese üldmõiste, mis hõlmab autoriõiguslikult kaitstavaid teoseid ja autoriõigusega kaasnevate õiguste eset (nt esitust).



Joonis 1. Keeleressursse katvad kahetasemelised õigused

3. Lepinguline ahel keeleressursside loomisest kasutamiseni

Õiguslikud aspektid omavad kõige otsesemat mõju keeleressursside levitamisele (litsentseerimisele). Keeleressursside litsentseerimise teemal on autorid avaldanud ühe eestikeelse üldise ülevaateartikli (Vider jt 2013) ja CLARIN-i tüüplepinguid konkreetsemalt analüüsiva artikli (Kelli jt 2015). Siin käsitleme töösuhtes loodud keeleressursside lepingulisi küsimusi ning CLARIN-i tüüplepinguid ja nende võimalikke parandusettepanekuid.

3.1. Töösuhtes loodud keeleressursid

Keeleressursse loovad füüsilised isikud, kuid loodud ressursse haldavad üldjuhul juriidilised isikud, näiteks Tartu Ülikool (TÜ), Tallinna Tehnikaülikool (TTÜ Küberneetika Instituut) ja Eesti Keele Instituut (EKI). Seetõttu on oluline, et õigused loodud ressurssidele läheksid üle juriidilistele isikutele. Autoriõiguse seadus reguleerib küll töösuhtel loodud autoriõiguse ja kaasnevate õiguste üleminekut, kuid sellega seonduvad järgmised probleemid:

- 1) tööülesanded peavad olema väga selgelt defineeritud (Kurisoo 2010), sest autoriõiguse seadus kasutab väljendit „oma otseste tööülesannete täitmise korras loodud teos” (§ 32). Seetõttu võivad tekkida vaidlused, mis on otsesed tööülesanded;
- 2) autoriõiguse ja kaasnevate õiguste esemed on allutatud erinevale kuuluvusrežiimile: töösuhtes loodud autoriõigusega kaitstava teose varalised õigused lähevad üle tööandjale (§ 32 lg 1). Arvutiprogrammide ja andmebaasile, mis olemuselt on samuti autoriõigusega kaitstavad teosed, saab tööandja ainulitsentsi (§ 32 lg 5). Õigused esitusele jäävad üldse esitajale, kui ei ole otseselt kokku lepitud teisiti (§ 67 lg 5). Autoriõiguse eelnõuga on plaan reguleerida töösuhtes loodud teoste ja kaasnevate õiguste objekti kuuluvust lähtudes ühtsetest alustest. Autoriõiguse eelnõu sätestab järgmise regulatsiooni (§ 23 lg 1):

“Töölepingu alusel oma tööülesannete täitmise korras loodud teose autoril tekib autoriõigus sellele teosele, kuid autori varalised õigused teose kasutamiseks tööülesannetega ettenähtud eesmärgil lähevad üle

tööandjale, kui autori ja tööandja vahelise kokkuleppega ei ole ette nähtud teisiti”.

Sama lähenemine on ka esitajate puhul (§ 29 lg 2);

- 3) töösuhtes lähevad üle varalised õigused, kuid mittevaralised õigused, mille hulka kuulub õigus teose puutumatusel⁸ ja õigus teost avalikustada, on isiklikud õigused (§ 12), mille teostamise osas tuleb eraldi kokku leppida. Autoriõiguse ja autoriõigusega kaasnevate õiguste seaduse eelnõu annab siinkohal järgmise juhtnööri (§ 10 lg 3):

“Autor võib selgesõnalise kirjaliku tahteavaldusega tähtajaliselt või tähtajatult loobuda oma isiklike õiguste teostamisest või anda nõusoleku oma isiklike õiguste teostamiseks teise isiku poolt. Nimetatud tahteavaldust ei saa tagasi võtta”.

Omaette probleemide ring seondub veel keeleressursside loomisega teenuste osutamise lepingute (nt töövõtu- ja käsunduslepingu) alusel. Kuna võlaõigusseadus ei reguleeri otseselt töövõtu- või käsundi täitmise tulemusena loodud intellektuaalse omandi (IO) kuuluvust, siis see võib tekitada vaidlusi. Võlaõigusseaduse kommentaarides avaldatakse arvamust, et töövõtu alusel loodud teose varalised õigused peaks minema üle tellijale, kuid samas mööndakse, et regulatsioon on ebaselge (Varul jt 2009: 43). Selguse huvides on oluline nii töösuhtes kui ka võlaõiguslike lepingute alusel loodud keeleressursside kuuluvuses kokku leppida. Lisaks peaks olema selge kokkulepe ka ressursside muutmise, täiendamise, ümbertöötamise, teiste ressurssidega liitmise (st autori isiklike õiguste) osas.

3.2. Eesti Keeleressursside Keskus kui lepingupartner

Tartu Ülikool, Tallinna Tehnikaülikooli Küberneetika Instituut ja Eesti Keele Instituut on loonud konsortsiumlepingu alusel Eesti Keeleressursside Keskuse (EKRK), mille ülesandeks on teha eesti keele digitaalsed ressursid ja tehnoloogiad kättesaadavaks. Ühtlasi on EKRK-l kohustus koguda kokku ja teha kättesaadavaks riikliku programmi “Eesti keeletehnoloogia” projektide tulemused. Selleks, et seda teha, peab EKRK-l olema mingi õiguslik alus saada digitaalsete keeleressursside suhtes õigused neid teoseid ise kasutada ja anda kasutada või teha muul viisil kättesaadavaks kolmandatele isikutele. Kasutusõigusi ja -tingimusi sätestatakse sel juhul deponeerimis- ja litsentsilepingutes.

Juriidilisest perspektiivist vaadates pole EKRK iseseisev juriidiline isik, vaid üksnes konsortsium- ehk seltsinguleping, millega kooskõlas lepingupartnerid tegutsevad (seltsingulepingu kohta vt Varul jt 2007: 635–727). Üldistatult saab öelda, et õigusi ja kohustusi saavad võtta üksnes konsortsiumi liikmed kui juriidilised isikud ja EKRK saab lepinguid sõlmida üksnes partnerite nimel.

Vastavalt EKRK konsortsiumipartnerite ning Haridus- ja Teadusministeeriumi vahelisele lepingule (CLARIN ERIC-u riikliku konsortsiumi leping) täidab EKRK ka CLARIN ERIC-u kui Euroopa teadusuuringute infrastruktuuri konsortsiumi Eesti riikliku konsortsiumi ülesandeid. Toodud aspektid on olulised, kui me hakkame vaatlema CLARIN-i tüüplepingute temaatikat.

⁸ Autoriõiguse seadus määratleb autori isiklikku õigust teose puutumatusel järgmiselt: “Teose autoril on õigus teha ise või lubada teha teistel isikutel teoses endas, tema pealkirjas (nimetuses) või autorinime tähistuses mis tahes muudatusi ning õigus vaidlustada ilma autori nõusolekuta tehtud muudatusi (õigus teose puutumatusel)” (§ 12 lg 1 punkt 3).

3.3. CLARIN-i tüüplepingud

CLARIN-i eesmärgiks on ühtlustada muuhulgas ka oma liikmete tüüplepingute raamistik, mille alusel deponeeritakse ressursid ja antakse neid kasutada. CLARIN-i tüüplepingute väljatöötamise algaasis Eesti arvestatavalt ei osalenud. Eesti on hakanud intensiivsemalt panustama CLARIN-i tüüplepingute täiustamisesse alates 2014. aastast. Mitmed CLARIN-i tüüplepingude puudujäägid on tulnud ilmsiks nende eesti keelde tõlkimise protsessis. Enne võimalike muudatuste lähemat analüüsi peatume CLARIN-i tüüplepingute väljatöötamise taustsüsteemil.

CLARIN-i tüüplepingute raamistikus on keeleressursid (nende juhtimise huvides) jagatud kolme kategooriasse:

- 1) avalikult kättesaadavad ressursid (ingl *Publicly Available Resources*, PUB) – neid võib levitada ilma geograafiliste või kasutusvaldkonna piiranguteta,
- 2) teadustöökasutatavad ressursid (ingl *Resources for Academic Use*, ACA) – neid võib kasutada üksnes teadustöö eesmärgil,
- 3) piiratud kasutusega ressursid (ingl *Resources for Restricted Use*, RES) – nende kasutamine on kitsendatud isikuandmete kaitsest tulenevalt, eetilistel või muudel kaalutlustel (Oksanen jt 2010). CLARIN-i kategooriate tagapõhja on põhjalikumalt käsitletud CLARIN-i mudellitsentsi lepingute aruandes (CLARIN-i aruanne 2010).

Sellise kolmikjaotuse üle on väideldud, on tehtud ka muid ettepanekuid. Näiteks Pawel Kamocki⁹ pakub välja sellise liigenduse: CLARIN OPEN ja CLARIN DAO (ingl *Data Analysis Only*) ning täiendav kategooria CLARIN *Public Domain*. See jaotus vajaks aga põhjalikku eelnevat uurimistööd, sest see mõjutaks olemasolevaid lepinguid (need peaks ümber kvalifitseerima) ning uute kategooriate ülevõtmist peavad CLARIN-i liikmed toetama.

CLARIN-i tüüplepingud on inglise keeles ning kasutusel näidistena CLARIN-i võrgustiku liikmete jaoks. Eestis on CLARIN-i tüüplepingud (deponeerimislepingud ja kasutustingimused) tõlgitud eesti keelde, kohaldava õigusena on sätestatud Eesti õigus ning vaidluse lahendamise kohana Eesti kohus. Lisaks on eestikeelsetes tüüplepingu versioonides parandatud algsetes CLARIN-i tüüplepingutes esinenud vead. Eesti keelde tõlgitud ja toimetatud CLARIN-i tüüplepingud ei ole veel kasutuses, sest enne on vaja täiendada algseid CLARIN-i tüüplepinguid. Eesti eesmärk on tegutseda kooskõlas ülejäänud CLARIN-i liikmetega. Kui iga liige lähtub oma standardist, siis ei ole CLARIN-il eriti mõtet.

Üldistatult võib öelda, et CLARIN-i keeleressursi kategooria tuleneb ressursi loomiseks kasutatud sisendmaterjalist ja selle kasutamise alusest. See tingib ka konkreetse tüüplepingu kasutamise. Et keeleressursside loomine Eestis tugineb suures osas autoriõigusest tulenevale teadustöö erandile, siis kuulub Eestis loodud keeleressurss sageli ACA kategooria alla. Keeleressurss läheb PUB kategooria alla ja on vabalt levitatav siis, kui lähtematerjalile laienevate õiguste omaja (nt autor, kirjastaja jne) on andnud selleks loa, keeleressurss on loodud autoriõiguslikult mittekaitstavast materjalist (teosed, mille kaitsetähtaeg on läbi; mittekaitstavad tekstid nagu seadused, kohtulahendid või rahvalooming), autoriõiguslik materjal ei kajastu keeleressurssis või siis, kui keeleressurss koosneb lühikestest tsitaatidest.

⁹ Autorid suhtlesid CLARIN ERIC-u õiguskomisjoni liikme Pawel Kamockiga e-kirja teel (1.10.2015).

Kui keeleressurss sisaldab tundlikke isikuandmeid, võib see kuuluda RES kategooriasse, välja arvatud juhul, kui ei ole just andmesubjekti selget ja teadvat nõusolekut isikuandmete avalikustamiseks.

Ressursi levitamine saab alguse selle üleandmisest või deponeerimisest CLARIN-i keskuses (<https://keeleressurssid.ee/et/clarin>). EKRK konsortsiumipartneritele (TÜ, TTÜ, EKI) kuuluvate ressursside osas ei ole vaja sõlmida deponeerimislepinguid. Nende jaoks on oluline, et töölepingute intellektuaalse omandi kokkulepped oleksid asjakohaselt sõnastatud ning eksisteeriks õiguste üleandmise aktid. CLARIN-i deponeerimise tüüplepinguid (DELA) saab kasutada suhetes kolmandate isikutega, kes ei loonud ressursse TÜ, TTÜ või EKI töötajadena, aga tahavad ressursse levitada CLARIN-i võrgustikus. Kuna EKRK on konsortsium, sõlmivad deponeerimislepinguid TÜ, TTÜ või EKI vastavalt oma tegevusvaldkonnale. CLARIN-i deponeerimislepingud on sarnaselt ressurssidele jagatud kolme kategooriasse:

- 1) avatud keeleressursside CLARIN-i teenusesse lisamise leping (DELA-PUB 1.1);
- 2) teadustöök kasutatavate keeleressursside CLARIN-i teenusesse lisamise leping (DELA-ACA 1.1);
- 3) isikliku kasutamise eesmärgidel keeleressursside CLARIN-i teenusesse lisamise leping (DELA-RES 1.1).

Kõigi kolme deponeerimislepingu regulatsioonid kattuvad suures osas. Lepingu sisuks on keeleressursside üleandmine repositooriumile, mida haldab CLARIN-i volitatud keskus (lõppastmes TÜ, TTÜ või EKI) ning loa (litsentsi) andmine ressursside levitamiseks. DELA-ga reguleeritakse ka ressursi üleandmise tehnilisi küsimusi, hooldust ja ajakohastamist, lepingu kehtivust, kohaldatavat õigust ja muid küsimusi. Oluline säte on hoiuleandja (ressursi looja) vastutuse võtmine selle eest, et talle kuuluvad õigused ressursile ning kui ressursi levitamine rikub kolmanda isiku õigusi, siis kohustub hoiuleandja hüvitama CLARIN-i keskusele võimalikud kahjud.¹⁰ Uuendamata DELA versioonides ei ole see põhimõte kõige paremini sõnastatud ning vajab ümbersõnastamist. Täiendavalt oli vaja täpsustada DELA-des kasutatavat terminoloogiat, muuta sõnastust selgemaks jms. Ka DELA-de praegused versioonid ei ole lõplikud ning neisse tuleb suhtuda kui üldisesse käitumisjuhisesse, milles CLARIN-i liikmed on kokku leppinud.

DELA-de jagunemine kolme põhikategooriasse (PUB, ACA, RES) tuleneb intellektuaalse omandi sätetest ehk siis sellest, mis tingimustel võib deponeeritud ressursi levitada. Avatud keeleressursside CLARIN-i teenusesse lisamise lepingu kohaselt (punkt 7) teeb hoiuleandja ressursi kättesaadavaks ühe või mitme järgneva litsentsi alusel: *Creative Commons ZERO*, *Public Domain* (autoriõigusest loobutakse), *Princeton Wordnet*, CLARIN PUB. Täiendavad õigused seoses ressursiga võivad olla kirjalikult eraldi kokku lepitud. Teadustöök ja isiklikeks eesmärkideks kasutatavate keeleressursside CLARIN-i teenusesse lisamise lepingutega nähakse ette täiendavaid piiranguid ressursi kasutamisel ja levitamisel.

Keeleressursi loojate (hoiuleandjate) ja ka teiste abistamiseks on välja töötatud CLARIN-i litsentsi kategooria kalkulaator (CLARIN-i kalkulaator), mis küsimustele antud vastuste pinnalt kuvab kasutustingimused ning pakub sobiva tüüplitsentsi või annab soovitusel eraldi kasutusleping luua. Küsimused ja tingimused kalkulaatoris

¹⁰ Kolmandale isikule võib tekitada kahju ka nt ressursi ebatäpsus (mõiste on valesti defineeritud) ning sellele tuginedes mõni toote valesti koostatud kasutusõpetus vms. Seetõttu on oluline, et ressursside kasutamisel nähakse ette tingimus, et ressursside kasutamine toimub omal vastutusel. Ressursse pakutakse sellistena, nagu nad on (ingl *as is*).

on jagatud järgmiselt: 1) identifitseerimise ja juurdepääsu tingimused (ingl *Identification and access conditions*); 2) üldised kasutustingimused (ingl *General use conditions*); 3) levitustingimused (ingl *Distribution conditions*); 4) muud tingimused (ingl *Other conditions*). Sisuliselt vastavad need deponeerimislepingute ja kasutustingimuste vastavale alaosalale.

Probleeme on mõnikord tekkinud tervikteose ja teose osade (fragmentide) litsentsimise eristamisega. Autoriõiguslikult kaitstavad fragmendid on algeosest tuletatud teosed ning neile laieneb sama litsents, mis algeosele, kui ei ole teisiti kokku lepitud. Mõnel juhul võib algeose õiguste omaja lubada tuletatud fragmentide (lausetes või lõikudes) levitamist, kasutades näiteks korpuste otsingumootorit nagu KORP (vt <http://korp.keeleressursid.ee/> või <https://korp.csc.fi>), kuid samas ei ole ta nõus tervikteose levitamise. Nii on näiteks EK RK-i käes tekstikorpused kui tervikteos, aga kasutaja seda tervikuna endale alla laadida ei saa. KORP-i kaudu sõna- vms päringut esitades saab kasutaja ainult talle huvi pakkuvale keelematerjalile lause kaupa aga hõlpsasti ligi. Sellega seoses kerkib praktiline küsimus, kuidas selline põhimõte sõnastada DELA-s ja ressursi metaandmete kirjelduses, ilma et see tekitaks segadust lõppkasutaja õiguste osas.

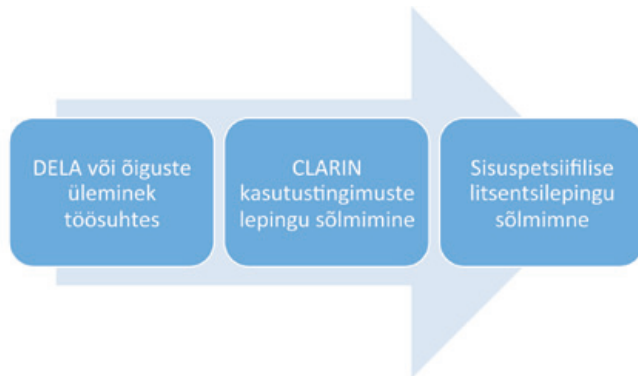
Kui keeleressurssi XXX ei või ühelgi tingimusel levitada, siis on sellel litsents, mis ei lähe isegi RES kategooria alla. Praktikast ei pruugi see ressurss olla kellelegi kättesaadav peale otsingumootori pakkuja. Samas otsingutulemused, mis koosnevad fragmentidest (lausetest ja lõikudest), võivad kuuluda PUB kategooria alla ja olla kõigile kättesaadavad. Ressursi XXX kirjeldamine PUB ressursina on aga eksitav, sest võib tekitada mulje, et kasutajal on võimalik kogu ressursi alla laadida. Lahendusena võiks näha DELA-s selgelt ette, et XXX ressursist tuletatud fragmente võib levitada. Selliseid tuletatud fragmente võib nimetada “XXX väljavõteteks”. Isegi kui XXX ressursi tervikuna ei ole lubatud levitada, siis otsingutulemused (tuletatud teoseid ehk XXX väljavõtteid) võib levitada kui PUB kategooria ressursse.

Lisaks tuleb lahendada veel üks tehniline küsimus. Sellistele tuletatud teostele nagu XXX väljavõtted ei saa tingimata anda püsidentifikaatorit (ingl *persistent identifier*, PID), sest uus tuletatud teos tekib iga kord, kui keegi teeb uue tulemust andva otsingu. Teoreetiliselt võiks iga uus otsingutulemus saada püsidentifikaatori. Samas tuleb tagada, et identse otsingutulemused saavad identse püsidentifikaatori. Metaandmete registrid nagu CLARIN-VLO ja META-SHARE võimaldavad registreerida ressursse (nagu XXX väljavõtted) ilma püsidentifikaatorita PUB kategoorias, vältides vastuolo registreeritud püsidentifikaatoriga XXX tervikressursiga.

Kui keeleressursid on läinud üle EK RK partneritele (TÜ, TTÜ ja EKI) kas töösuhte alusel või vastava (PUB, ACA, RES) deponeerimislepingu sõlmimise teel, siis on võimalik alustada nende levitamist.

Järgmine samm on ressursside kasutajatele kättesaadavaks tegemine. Enne kui kasutaja saab ressurssidele juurdepääsu, peab ta nõustuma kasutustingimustega. CLARIN-is on sel eesmärgil välja töötatud tüüptingimused (CLARIN-i teenusetingimused), mille punkt 3.6 näeb ette identifitseerimise ja juurdepääsu tingimused, üldised kasutustingimused, levitustingimused ja muud olulised tingimused. Täiendavalt võib olla vajalik sõlmida iga ressursi kasutamiseks spetsiifiline litsentsileping. Tähtis on ka arvestada, et keeleressursside kättesaadavaks tegemise ulatus sõltub TÜ, TTÜ, EKI või mõnele teisele kättesaadavaks tegijale kuuluvatest õigustest. Kui ressurss on loodud teadustöö erandile tuginedes, siis ei saa selle levitamisel anda

lõppkasutajale laiemaid õigusi (näiteks lubada ärilist kasutust). Deponeeritud ressursi puhul peavad lõppkasutajale antavad õigused sisuliselt peegeldama DELA tingimusi. See tähendab, et kui DELA-s nõutakse keeleressursside deponeerimisel uurimisplaani olemasolu, siis tuleb ka kasutustingimustes taoline nõue sätestada. Joonis 2 illustreerib lepingulist ahelat alates keeleressursside õiguste üleminekust kuni nende kasutamiseks spetsiifilise litsentsilepingu sõlmimiseni.



Joonis 2. Keeleressursside kättesaadavaks tegemise lepinguline raamistik

CLARIN-i kasutustingimused on keeleressurssidele juurdepääsu üldine raamistik. EKRR-i loonud TÜ, TTÜ ja EKI selgeks huviks on oma vastutuse piiramine seoses keeleressursside kättesaadavaks tegemisega. Seetõttu on vastutust piirav klausel (ingl *Disclaimer of Warranties and Limitation of Liability*) ka üks olulisemaid asju kättesaadavaks tegemise tingimuste kõrval. CLARIN-i teenusetingimuste eelmise versiooni vastutust piirava klausli (punkt 5) sõnastus ei olnud kõige õnnestunud ning seda tuli muuta. Oluline on näha ette, et ressursse pakutakse olemasolevas seisus (ingl *as is*), sellisena, nagu see on kättesaadav (ingl *as available*) ning CLARIN ei vastuta mingite kahjude eest (5. Garantii mitteandmine ja vastutuse piiramine).

Näidisena ja regulatsiooni väljatöötamise lähtealusena on siinkohal kasutatud *Creative Commons*i punktis 5 ja Euroopa Liidu tarkvara vaba kasutuse litsentsi (EUPL) punktides 7 ja 8 olevat regulatsiooni.

4. Kokkuvõte

Artiklis uurisid autorid keeleressursside temaatikat läbi õigusliku prisma. Käsitluses lähtuti protsessipõhisest lähenemisest, mis algab keeleressursside loomisega ja lõppeb nende levitamisega. Keeleressursside uurimise kontseptuaalseks aluseks oli CLARIN-i võrgustiku tegevus oma liikmete õigusliku ja lepingulise raamistiku ühtlustamisel. Kõige selgemini on see leidnud väljenduse CLARIN-i tüüplepingute väljatöötamisel ja siseriiklikul kohaldamisel.

Kogu protsess algab keeleressursside loomisest ning sellest sõltuvad ka keeleressursside levitamise võimalused. Keeleressurss on autoriõiguslikult kaitstav iseisva teosena (harilikult andmebaasina), kuid selle levitamiseks ei pruugi piisata keeleressursside looja poolt antavatest õigustest. Probleem seisneb selles, et sageli sisaldab keeleressurss teiste isikute poolt loodud teoseid, esitusi või isikuandmeid.

Kui keeleressurss on loodud teose vaba kasutamise teadustöö erandile tuginedes, siis ei saa ka seda teost levitada või edasi anda näiteks äriliste rakenduste loomiseks. Keeleressurssi iseloomustavad seega kahetasemelised õigused: keeleressursi looja ja sisendmaterjali omaja õigused.

Pärast keeleressursi loomist on oluline tagada selle üleandmine. Kui keeleressurss on loodud töösuhtes, siis lähevad õigused sellele üle seaduse alusel. Õiguselguse huvides tuleks õiguste üleminekus selgelt kokku leppida. Kuna keeleressursi loojal kui autoril on isiklikud õigused, mille hulka kuulub ka õigus teose puutumatusele (õigus teost muuta), siis tuleb ka selle õiguse teostamises kokku leppida. Keeleressursside loojatega, kes ei kuulu Eesti Keeleressursside Keskuse (EKRK) loonud konsortsiumi partnerite (TÜ, TTÜ, EKI) hulka, tuleb sõlmida deponeerimisleping.

CLARIN-i tüüplepingute järgi jagunevad keeleressursid kolme kategooriasse: PUB (avalik), ACA (akadeemiline) ja RES (piiratud). Sama jaotus on ka lepingute kategooriatel, kaasa arvatud CLARIN-i deponeerimislepingutel (DELA). Deponeerimislepinguga määratletakse keeleressursi levitamise tingimused. DELA-ga annab ressursi looja kinnituse, et ta on ressursi loonud ning vastutab ressursi levitamisest tuleneva võimaliku kahju eest.

Sõltumata sellest, kas ressurss on läinud üle töösuhtes või deponeerimislepingu alusel, on kasutaja jaoks ressursile juurdepääsu saamise eelduseks CLARIN-i teenusetingimustega nõustumine. CLARIN-i kasutustingimuste eesmärgiks on näha ette kasutamise üldine raamistik, kuid ka piirata ressursi kättesaadavaks tegija vastutust maksimaalses ulatuses. Lisaks CLARIN-i üldistele teenusetingimustele peab kasutaja nõus olema veel konkreetse keeleressursi litsentsitingimustega.

CLARIN-i lepinguline raamistik kujutab endast dünaamilist töövahendit, mida arendatakse ja reformitakse pidevalt. Juba praegu on alusdokumentides tuvastatud rida puudusi, mis on kõrvaldatud nii inglisi- kui ka eestikeelsetes versioonides.

Viidatud kirjandus

AutÕS = Autoriõiguse seadus. RT I 1992, 49, 615 ... RT I, 29.10.2014, 2.

Autoriõiguse eelnõu = Autoriõiguse ja autoriõigusega kaasnevate õiguste seaduse eelnõu.

Versioon: 27.1.2015. www.just.ee/sites/www.just.ee/files/autorioiguse_seaduse_eelnou_27.01.2015.pdf (10.12.2015).

Autoriõiguse eelnõu seletuskiri = Seletuskiri autoriõiguse ja autoriõigusega kaasnevate õiguste seaduse eelnõu juurde. Versioon: 21.7.2014. <https://ajaveeb.just.ee/intellektuaalne-omand/wp-content/uploads/2014/08/Autori%C3%B5iguse-seletuskiri-21-7-2014.pdf> (10.12.2015).

Avalik konsultatsioon 2014. (Report on the responses to the Public Consultation on the Review of the EU Copyright Rules. July 2014.) http://ec.europa.eu/internal_market/consultations/2013/copyright-rules/docs/contributions/consultation-report_en.pdf (10.12.2015).

Berni konventsioon = Berni kirjandus- ja kunstiteoste kaitse konventsioon. RT II 1994, 16, 49.

CLARIN = Common Language Resources and Technology Infrastructure. <http://clarin.eu/content/about-clarin> (24.9.2015).

CLARINi aruanne 2010. (CLARIN. A report including Model Licensing Templates and Authorization and Authentication Scheme (CLARIN-2010-12)). <http://www-sk.let.uu.nl/u/D7S-2.1.pdf> (8.12.2015).

CLARIN ERICu riikliku konsortsiumi leping. <http://193.40.113.42:8088/AT/index.php/dokumentiregister/2012/details/4/1227/2012-kuup%C3%A4ev-sissetulemise-v%C3%A4ljaminemise-from-01-jaanuar-2012-to-31-detsember-2012-2012-07-23?start=120> (10.12.2015).

CLARINi kalkulaator. <http://www.helsinki.fi/finclarin/calculator/ClarLicenseCategory.html> (10.12.2015).

CLARINi teenusetingimused = CLARIN Terms of Service (ToS) 1.1. <https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/ClarTermsOfServiceV11> (10.12.2015).

CLARIN PUB = CLARIN-EULA-PUB-v1.0. <https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/ClarEULA#pub> (10.12.2015).

CLARIN-VLO. <https://vlo.clarin.eu/> (3.10.2015).

Cornish, William; Llewelyn, David 2007. Intellectual Property: Patents, Copyright, Trade Marks and Allied Rights. Sixth Edition. London: Sweet & Maxwell.

Correa, Carlos María 2007. Trade Related Aspects of Intellectual Property Rights: A Commentary on the TRIPS Agreement. Oxford: Oxford University Press.

Creative Commons = Creative Commons. About The Licenses. <http://creativecommons.org/licenses/> (10.12.2015).

Creative Commons ZERO. <http://creativecommons.org/publicdomain/zero/1.0/legalcode.txt> (10.12.2015).

DELA-ACA v1.1 = CLARIN ACA Deposition License Agreement Template 1.1. <https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/ClarDelaV11?ACA=1> (10.12.2015).

DELA-PUB 1.1 = CLARIN PUB Deposition License Agreement Template 1.1. <https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/ClarDelaV11?PUB=1> (10.12.2015).

DELA-RES 1.1 = CLARIN RES Deposition License Agreement Template 1.1. <https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/ClarDelaV11?RES=1> (10.12.2015).

Eesti keeletehnoloogia = Riiklik programm "Eesti keeletehnoloogia 2011–2017". <https://www.keeletehnoloogia.ee/et> (10.12.2015).

EKRK = Eesti Keeleressursside Keskus. <https://keeleressursid.ee/et/eesti-keeleressursside-keskus> (10.12.2015).

ELRC = European Language Resource Coordination. <http://www.lr-coordination.eu> (8.12.2015).

Ettepanek infoühiskonna direktiivi rakendamiseks = Euroopa parlamendi resolutsiooni ettepanek Euroopa Parlamendi ja nõukogu 22. mai 2001. aasta direktiivi 2001/29/EÜ (autoriõiguse ja sellega kaasnevate õiguste teatavate aspektide ühtlustamise kohta infoühiskonnas) rakendamise kohta. 2014/2256(INI). <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+REPORT+A8-2015-0209+0+DOC+XML+Vo//ET> (10.12.2015).

EUPL = Euroopa Liidu tarkvara vaba kasutuse litsents. https://joinup.ec.europa.eu/sites/default/files/eupl1.1-licence-et_o.pdf (10.12.2015).

Euroopa autoriõiguse koodeks 2010 (European Copyright Code. 2010) www.copyrightcode.eu (8.12.2015).

Euroopa digitaalse ühtse turu strateegia 2015 = Komisjoni teatis euroopa parlamendile, nõukogule, euroopa majandus- ja sotsiaalkomiteele ning regioonide komiteele Euroopa digitaalse ühtse turu strateegia. COM (2015) 192 final. Brüssel, 6.5.2015. <http://eur-lex.europa.eu/legal-content/ET/TXT/?qid=1449739404979&uri=CELEX:52015DC0192> (10.12.2015).

Gibert, Benjamin 2015. The 2015 Intellectual Property and Economic Growth Index. Measuring the Impact of Exceptions and Limitations in Copyright on Growth, Jobs and Prosperity. Revised May 2015. The Lisbon Council, 2015. <http://www.innovationeconomics.net/component/attachments/attachments.html?id=263&task=view> (10.12.2015).

Gorman, Robert A.; Ginsburg, Jane C. 2006. Copyright Cases and Materials. Seventh edition. Foundation Press.

- Hohmann, Iris; Kullas, Mathias 2015. Exceptions to Copyright. Should optional exceptions be made mandatory? Centrum für Europäische Politik. http://www.cep.eu/Studien/cepInput_Copyright/cepInput_Exceptions_to_copyright.pdf (10.12.2015).
- IKS = Isikuandmete kaitse seadus. – RT I 2007, 24, 127 ... RT I, 29.06.2014, 109.
- Infoühiskonna direktiiv = Euroopa Parlamendi ja Nõukogu direktiiv 2001/29/EÜ, 22. mai 2001, autoriõiguse ja sellega kaasnevate õiguste teatavate aspektide ühtlustamise kohta infoühiskonnas. EÜT L 167, 22.6.2001, 10–19, eestikeelne eriväljaanne: ptk 17, kd 001, 230–239.
- Jents, Liina; Kelli, Aleksei 2014. Legal Aspects of Processing Personal Data in Development and Use of Digital Language Resources: the Estonian Perspective. *Jurisprudence*, 21 (1), 164–184. <https://www.mruni.eu/upload/iblock/da7/JUR-14-21-1-o8.pdf> (8.12.2015).
- Kelli, Aleksei; Tavast, Arvi; Pisuke, Heiki 2012. Copyright and Constitutional Aspects of Digital Language Resources: The Estonian Approach. – *Juridica International* 19, 40–48. http://www.juridicainternational.eu/public/pdf/ji_2012_1_40.pdf (8.12.2015).
- Kelli, Aleksei; Vider, Kadri; Lindén, Krister 2015. The Regulatory and Contractual Framework as an Integral Part of the CLARIN Infrastructure: an Estonian and Finnish Perspectives. – CLARIN Annual Conference 2015. Book of Abstracts: CLARIN Annual Conference, Wrocław, Poland, October 14–16, 2015, 32–36. <http://www.clarin.eu/sites/default/files/book%20of%20abstracts%202015.pdf> (7.12.2015).
- Kelli, Aleksei 2015. The conceptual bases for codifying Estonia's IP law and the main legislative changes: From the comparative approach to embedding drafted law into the socio-economic context. – *International Comparative Jurisprudence*, 1 (1), 44–54. <http://dx.doi.org/10.1016/j.icj.2015.10.004>
- Kelli, Aleksei; Hoffmann, Thomas; Pisuke, Heiki; Kull, Irene; Jents, Liina; Ginter, Carri 2014. The Exercise of Moral Rights by Non-Authors. – *Journal of the University of Latvia "Law"*, 6, 108–125.
- Kuriso, Kairi 2010. Mida teha, et töötaja oma autoriõigusega minema ei jalutaks? <http://www.director.ee/mida-teha-et-ttaja-oma-autoriigusega-minema-ei-jalutaks/> (7.12.2015).
- LT-Innovate Summit 2015. <http://www.lt-innovate.org/summit> (8.12.2015).
- META-SHARE = EKRK META-SHARE keeleressursside register. <https://metashare.ut.ee/> (10.12.2015).
- Nemvalts, Kärt; Seppel, Toomas; Kelli, Aleksei; Vasamäe, Elise; Elunurm, Silja; Adamsoo, Reet 2012. Autoriõiguse ja autoriõigusega kaasnevate õiguste probleemide koondkaardistus (17. juuli 2012). http://www.just.ee/sites/www.just.ee/files/elfinder/article_files/autorioiguse_koondkaardistus_o.pdf (7.12.2015).
- Oksanen, Ville; Lindén, Krister; Westerlund, Hanna 2010. Laundry Symbols and License Management: Practical Considerations for the Distribution of LRs based on experiences from CLARIN. <http://hdl.handle.net/10138/29359> (27.9.2015).
- Pisuke, Heiki 2006. Autoriõiguse alused. Tallinn: KPMS & Partnerid.
- Princeton Wordnet. <http://wordnet.princeton.edu/wordnet/license/> (10.12.2015).
- Ricketson, Sam; Ginsburg, Jane C. 2006. *International Copyright And Neighbouring Rights: The Berne Convention and Beyond*. Second edition. Vol. I. Oxford: Oxford University Press.
- Riga Summit 2015 on the Multilingual Digital Single Market. <http://www.rigasummit2015.eu> (8.12.2015).
- Tavast, Arvi; Pisuke, Heiki; Kelli, Aleksei 2013. Õiguslikud väljakutsed ja võimalikud lahendused keeleressursside arendamisel. – *Eesti Rakenduslingvistika Ühingu aastaraamat*, 9, 317–332. <http://dx.doi.org/10.5128/ERYa9.20>
- TRIPS leping = Intellektuaalomandi õiguste kaubandusaspektide leping. RT II 1999, 22, 123.
- Varul, Paul, Kull, Irene; Kõve, Villu; Käerdi, Martin (Toim.) 2009. *Võlaõigusseadus III. Kommenteeritud väljaanne*. Tallinn: Kirjastus Juura.

Varul, Paul, Kull, Irene; Kõve, Villu; Käerdi, Martin (Toim.) 2007. Võlaõigusseadus II. Kommenteeritud väljaanne. Tallinn: Kirjastus Juura.

Vider, Kadri; Pisuke, Heiki; Mõtsküla, Peeter P.; Tuulik, Triin; Kelli, Aleksei 2013. Keele-ressursside litsentsimise õiguslikke küsimusi. – Eesti Rakenduslingvistika Ühingu aastaraamat, 9, 333–347. <http://dx.doi.org/10.5128/ERYa9.21>

VÕS = Võlaõigusseadus. – RT I 2001, 81, 487 ... RT I, 23.03.2015, 3.

Walter, Michel; von Lewinski, Silke (Ed.) 2010. European Copyright Law. A Commentary. Oxford: Oxford University Press.

Aleksei Kelli on intellektuaalse omandi dotsent (Tartu Ülikool) ning intellektuaalse omandi õiguse kodifitseerimise töögrupi juht (Justiitsministeerium). Ta tegeleb digitaalsete keeleressursside õiguslike küsimustega Tartu Ülikoolis ning Eesti Keele Instituudis ning ta on intellektuaalse omandi ekspert CLARIN-i õigusasjade komitees.

Kaarli pst 3, 10119 Tallinn, Estonia
aleksei.kelli@ut.ee

Kadri Vider on Eesti Keeleressursside Keskuse (EKRK) tegevjuht, CLARIN ERIC-u riiklik koordinaator Eestis ja Tartu Ülikooli arvutiteaduse instituudi keeletehnoloogia teadur. EKRK-i ülesandeid on mh riiklike keeletehnoloogia teadus- ja arendusprogrammide tulemusena loodud ressursside ja tarkvara haldamine, kättesaadavaks tegemine ning litsentsimine või sellealane konsulteerimine.

J. Liivi 2, 50409 Tartu, Estonia
kadri.vider@ut.ee

Heiki Pisuke on Euroopa Liidu Nõukogu peasekretariaadi osakonnajuhataja (Brüssel) ja Tartu Ülikooli külalisprofessor. Ta oli esimene TÜ intellektuaalse omandi õiguse professor. Professor Pisuke oli 1991–1994 kultuuriministeeriumi autoriõiguse töögrupi juht, mis töötas välja uue autoriõiguse seaduse ja valmistas ette Eesti ühinemise rahvusvaheliste konventsioonidega. Ta on olnud mitmete rahvusvaheliste organisatsioonide ja töögruppide liige. Artiklis väljendatakse autori isiklike arvamusi ja see ei kajasta EL-i nõukogu ning teiste EL-i institutsioonide seisukohti.

Kaarli pst 3, 10119 Tallinn, Estonia
heiki.pisuke@ut.ee

Krister Lindén on Helsingi Ülikooli nüüdiskeelte osakonna juhtivteadur ja Soome CLARIN-i konsortsiumi FIN-CLARIN teadusdirektor, samuti CLARIN-i õigusasjade komitee esimees. Aastatel 2008–2011 juhtis ta ESFRI projekti CLARIN õigusteemade töörühma. Teadustöös on tema peamised huvid seotud soome ja mitmekeelsete wordnet'ide ja mitmekeelsete puudepankadega.

Unioninkatu 40, 00014 Helsingin yliopisto, Finland
krister.linden@helsinki.fi

REGULATORY FRAMEWORK DETERMINING THE DEVELOPMENT AND UTILIZATION OF DIGITAL LANGUAGE RESOURCES IN ESTONIA AND ITS COMPATIBILITY WITH CLARIN INFRASTRUCTURE

Aleksei Kelli¹, Kadri Vider^{1,2}, Heiki Pisuke^{1,3}, Krister Lindén⁴

University of Tartu¹, Center of Estonian Language Resources²,
Council of the European Union³, University of Helsinki⁴

The article focuses on legal issues relating to language resources. These issues are analyzed throughout the process of the creation of language resources and their subsequent distribution. Language resources can be defined as copyright protected databases, which are governed by two tiers of rights: 1) rights in material used to create the database and 2) rights in this database. The creation of language resources can be subject to two models: 1) the contract model and 2) the exception model. Both models have their advantages and disadvantages. The models are analyzed in the light of developments at the EU level.

Before language resources can be distributed, it is necessary to transfer copyright covering resources to a partner of the Center of Estonian Language Resources. The rights are transferred within employment or contractually. For this purpose, CLARIN deposition license agreement templates have been developed. Before end users can access language resources they have to accept CLARIN Terms of Service.

The scope of the article is not confined to the explanation of CLARIN agreement templates. The article also highlights the latest amendments to these templates.

Keywords: regulatory and contractual framework, copyright, categories of language resources, CLARIN agreement templates, liability